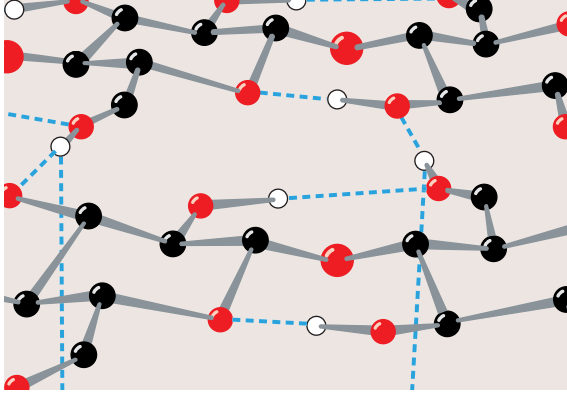


# Sugars and Polysaccharides

CHAPTER 11


1 **Monosaccharides**
- A. Classification
- B. Configurations and Conformations
- C. Sugar Derivatives

2 **Polysaccharides**
- A. Carbohydrate Analysis
- B. Disaccharides
- C. Structural Polysaccharides: Cellulose and Chitin
- D. Storage Polysaccharides: Starch and Glycogen
- E. Glycosaminoglycans

3 **Glycoproteins**
- A. Proteoglycans
- B. Bacterial Cell Walls
- C. Glycoprotein Structure and Function
- D. Glycomics


**Carbohydrates** or **saccharides** (Greek: *sakcharon,* sugar) are essential components of all living organisms and are, in fact, the most abundant class of biological molecules. The name carbohydrate, which literally means "carbon hydrate," stems from their chemical composition, which is roughly $(C \cdot H_2O)_n$, where $n \geq 3$. The basic units of carbohydrates are known as **monosaccharides.** Many of these compounds are synthesized from simpler substances in a process named **gluconeogenesis** (Section 23-1). Others (and ultimately nearly all biological molecules) are the products of **photosynthesis** (Section 24-3), the light-powered combination of $CO_2$ and $H_2O$ through which plants and certain bacteria form "carbon hydrates." The metabolic breakdown of monosaccharides (Chapters 17 and 21) provides much of the energy used to power biological processes. Monosaccharides are also principal components of nucleic acids (Section 5-1A), as well as important elements of complex lipids (Section 12-1D).

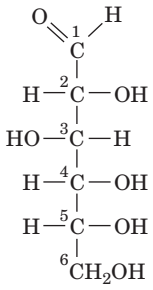
**Oligosaccharides** consist of a few covalently linked monosaccharide units. They are often associated with proteins **(glycoproteins)** and lipids **(glycolipids)** in which they have both structural and regulatory functions (glycoproteins and glycolipids are collectively called **glycoconjugates**). **Polysaccharides** consist of many covalently linked monosaccharide units and have molecular masses ranging well into the millions of daltons. They have indispensable structural functions in all types of organisms but are most conspicuous in plants because **cellulose,** their principal structural material, comprises up to 80% of their dry weight. Polysaccharides such as **starch** in plants and **glycogen** in animals serve as important nutritional reservoirs.

The elucidation of the structures and functions of carbohydrates has lagged well behind those of proteins and nucleic acids. This can be attributed to several factors. Carbohydrate compounds are often heterogeneous, both in size and in composition, which greatly complicates their physical and chemical characterization. They are not subject to the types of genetic analysis that have been invaluable in the study of proteins and nucleic acids because saccharide sequences are not genetically specified but are built up through the sequential actions of specific enzymes (Section 23-3B). Furthermore, it has been difficult to establish assays for the biological activities of polysaccharides because of their largely passive roles. Nevertheless, it is abundantly clear that carbohydrates are essential elements in many, if not most, biological processes.

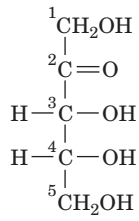
In this chapter, we explore the structures, chemistry, and, to a limited extent, the functions of carbohydrates, alone and in association with proteins. Glycolipid structures are considered in Section 12-1D. The biosynthesis of complex carbohydrates is discussed in Section 23-3.

## 1 MONOSACCHARIDES

Monosaccharides or **simple sugars** are aldehyde or ketone derivatives of straight-chain polyhydroxy alcohols containing at least three carbon atoms. Such substances, for example, **D-glucose** and **D-ribulose,** cannot be hydrolyzed to form simpler saccharides.



**D-Glucose**



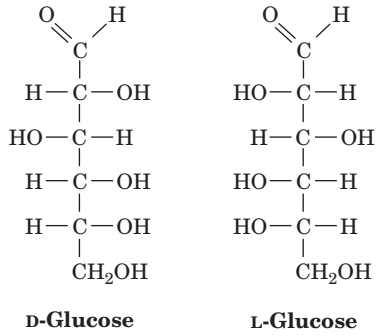
**D-Ribulose**

In this section, the structures of the monosaccharides and some of their biologically important derivatives are discussed.

## A. *Classification*

*Monosaccharides are classified according to the chemical nature of their carbonyl group and the number of their C atoms.* If the carbonyl group is an aldehyde, as in glucose, the sugar is an **aldose.** If the carbonyl group is a ketone, as in ribulose, the sugar is a **ketose.** The smallest monosaccharides, those with three carbon atoms, are **trioses.** Those with four, five, six, seven, etc., C atoms are, respectively, **tetroses, pentoses, hexoses, heptoses,** etc. These terms may be combined so that, for example, glucose is an **aldohexose,** whereas ribulose is a **ketopentose.**

Examination of D-glucose's molecular formula indicates that all but two of its six C atoms—C1 and C6—are chiral centers, so that D-glucose is one of $2^4 = 16$ stereoisomers that comprise all possible aldohexoses. In general, *n*-carbon aldoses have $2^{n-2}$ stereoisomers. The stereochemistry and names of the D-aldoses are presented in Fig. 11-1. Emil Fischer elucidated these configurations for the aldohexoses in 1896. According to the Fischer convention (Section 4-2B), *D sugars have the same absolute configuration at the asymmetric center farthest removed from their carbonyl group as does D-glyceraldehyde.* The L sugars, in accordance with this convention, are mirror images of their D counterparts, as is shown below in Fischer projection for glucose.



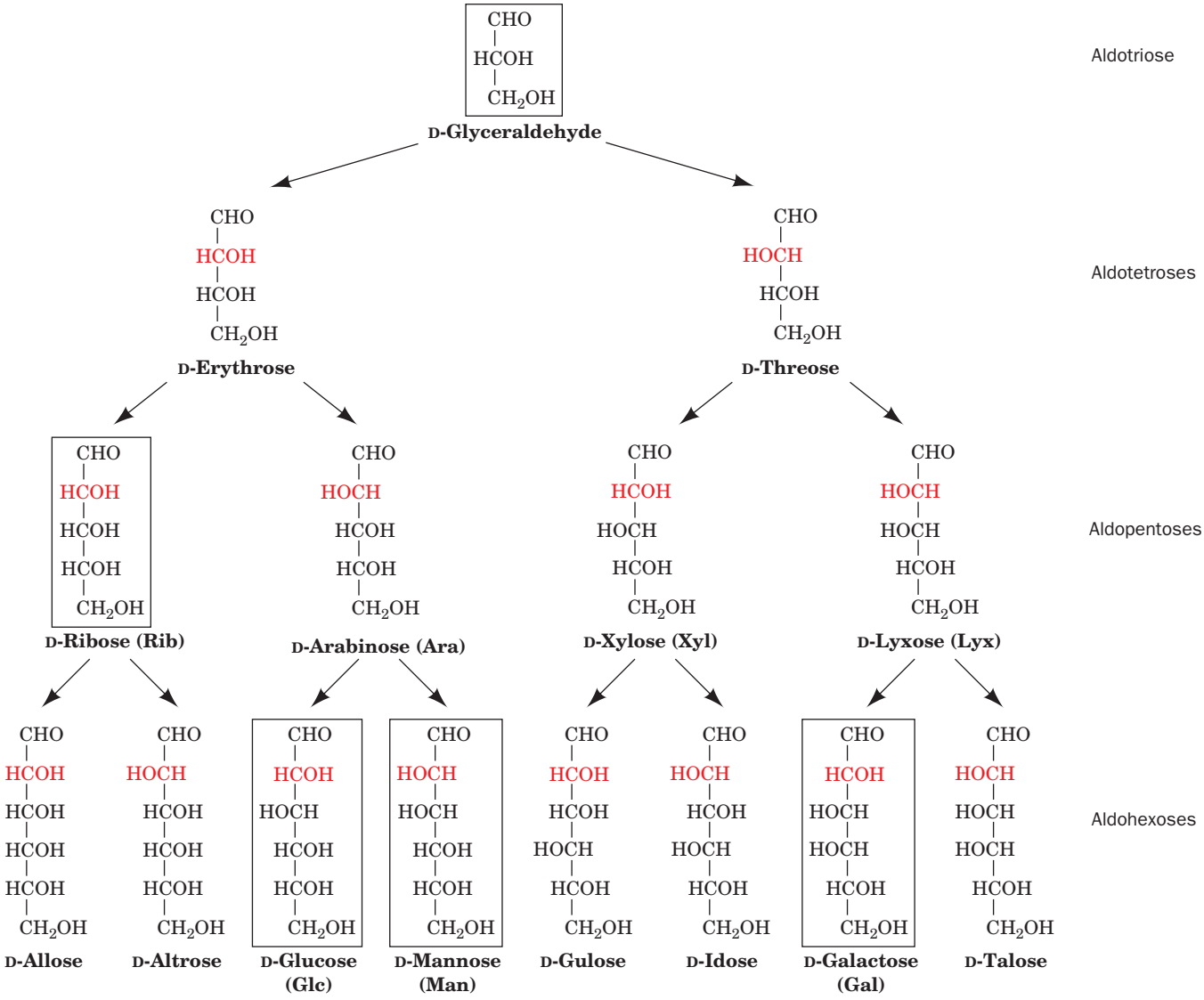




***Figure 11-1*** **The stereochemical relationships, shown in Fischer projection, among the D-aldoses with three to six carbon atoms.** The arrows indicate stereochemical relationships (not biosynthetic reactions). The configuration about C2 (*red*) distinguishes the members of each pair. The L- counterparts of these 15 sugars are their mirror images. The biologically most common aldoses are boxed.

Sugars that differ only by the configuration about one C atom are known as **epimers** of one another. Thus D-glucose and **D-mannose** are epimers with respect to C2, whereas D-glucose and **D-galactose** are epimers with respect to C4 (Fig. 11-1). However, D-mannose and D-galactose are not epimers of each other because they differ in configuration about two of their C atoms.

*D-Glucose is the only aldose that commonly occurs in nature as a monosaccharide.* However, it and several other monosaccharides including D-glyceraldehyde, D-ribose, D-mannose, and D-galactose are important components of larger biological molecules. L Sugars are biologically much less abundant than D sugars.

The position of their carbonyl group gives ketoses one less asymmetric center than their isomeric aldoses (e.g., compare D-fructose and D-glucose). *n*-Carbon ketoses therefore have $2^{n-3}$ stereoisomers. Those with their ketone function at C2 are the most common form (Fig. 11-2). Note that some of these ketoses are named by the insertion of *-ul-* before the suffix *-ose* in the name of the corresponding aldose; thus **D-xylulose** is the ketose corresponding to the aldose **D-xylose. Dihydroxyacetone, D-fructose,** D-ribulose, and D-xylulose are the biologically most prominent ketoses.

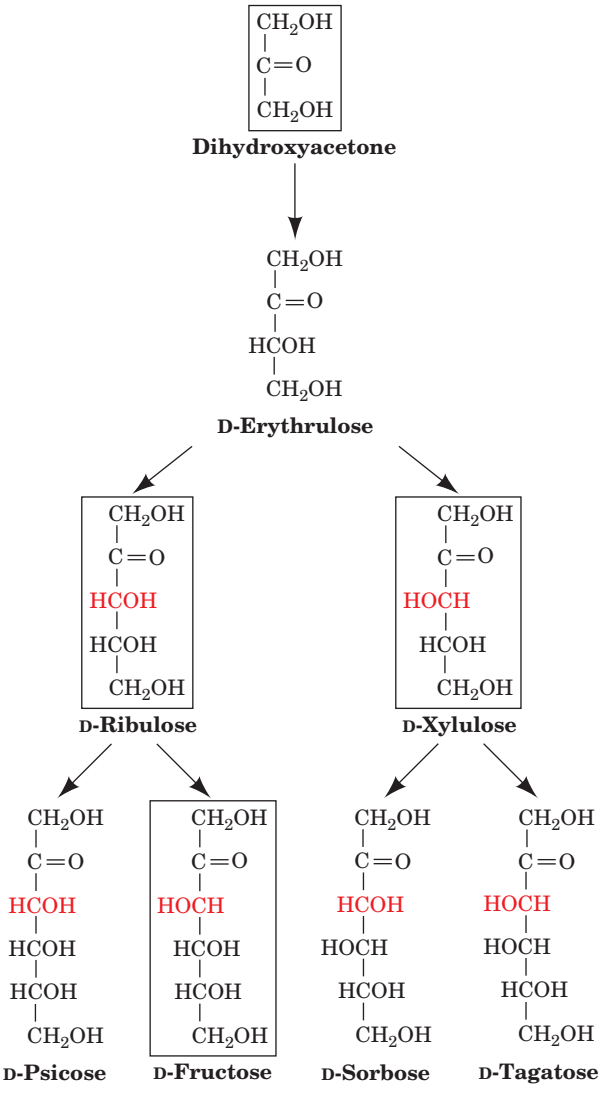
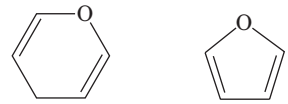



***Figure 11-2*** **The stereochemical relationships among the D-ketoses with three to six carbon atoms.** The configuration about C3 (*red*) distinguishes the members of each pair. The biologically most common ketoses are boxed.

## B. *Configurations and Conformations*

Alcohols react with the carbonyl groups of aldehydes and ketones to form **hemiacetals** and **hemiketals,** respectively (Fig. 11-3). The hydroxyl and either the aldehyde or the ketone functions of monosaccharides can likewise react intramolecularly to form cyclic hemiacetals and hemiketals (Fig. 11-4). The configurations of the substituents to each carbon atom of these sugar rings are conveniently represented by their **Haworth projection formulas.**

A sugar with a six-membered ring is known as a **pyranose** in analogy with **pyran,** the simplest compound containing such a ring. Similarly, sugars with five-membered rings are designated **furanoses** in analogy with **furan.**




The cyclic forms of glucose and fructose with six- and five-membered rings are therefore known as **glucopyranose** and **fructofuranose,** respectively.

### a. Cyclic Sugars Have Two Anomeric Forms

The Greek letters preceding the names in Fig. 11-4 still need to be explained. The cyclization of a monosaccharide renders the former carbonyl carbon asymmetric. The resulting pair of diastereomers are known as **anomers** and the hemiacetal or hemiketal carbon is referred to as the **anomeric** carbon. In the α anomer, the OH substituent to the anomeric carbon is on the opposite side of the sugar ring from the $CH_2OH$ group at the chiral center that

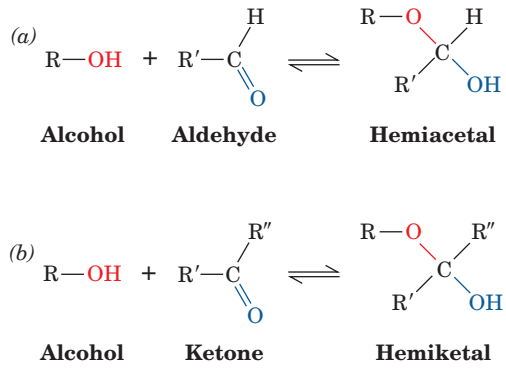



***Figure 11-3*** **The reactions of alcohols with (*a*) aldehydes to form hemiacetals and (*b*) ketones to form hemiketals.**

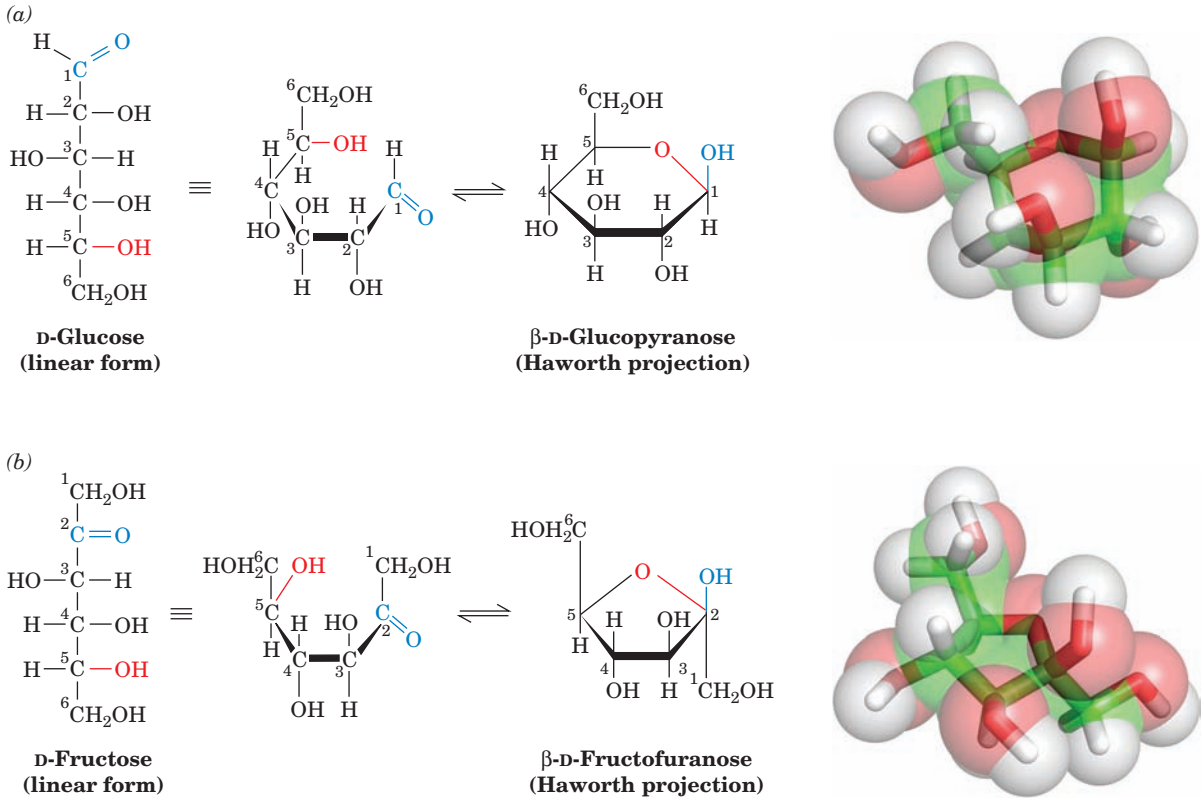



***Figure 11-4*** **Cyclization reactions for hexoses.** (*a*) D-Glucose in its linear form reacting to yield the cyclic hemiacetal β-D-glucopyranose and (*b*) D-fructose in its linear form reacting to yield the hemiketal β-D-fructofuranose. The cyclic sugars are shown both as Haworth projections and in stick form embedded in their semitransparent space-filling models with C green, H white, and O red.

designates the D or L configuration (C5 in hexoses). The other anomer is known as the β form (Fig. 11-5).

The two anomers of D-glucose, as any pair of diastereomers, have different physical and chemical properties. For example, the values of the specific optical rotation, $[\alpha]_D^{20}$, for α-D-glucose and β-D-glucose are, respectively, +112.2° and +18.7°. When either of these pure substances is dissolved in water, however, the specific optical rotation of the solution slowly changes until it reaches an equilibrium value of $[\alpha]_D^{20} = +52.7°$. This phenomenon is known as **mutarotation;** in glucose, it results from the formation of an equilibrium mixture consisting of 63.6% of the β anomer and 36.4% of the α anomer (the optical rotations of separate molecules in solution are independent of each other so that the optical rotation of a solution is the weighted average of the optical rotations of its components). The interconversion between these anomers occurs via the linear form of glucose (Fig. 11-5). Yet, since the linear forms of these monosaccharides are normally present in only minute amounts, these carbohydrates are

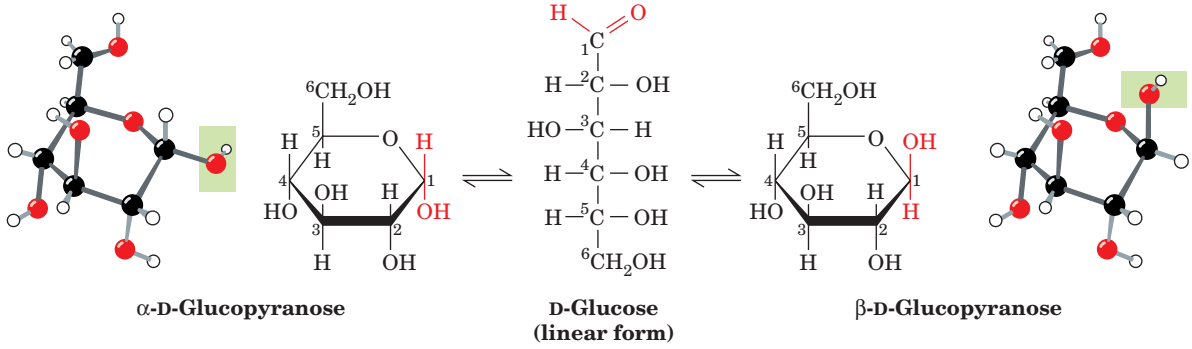



***Figure 11-5*** **The anomeric monosaccharides α-D-glucopyranose and β-D-glucopyranose, drawn as both Haworth projections and ball-and-stick models.** These pyranose sugars interconvert through the linear form of D-glucose and differ only by the configurations about their anomeric carbon atoms, C1.

See Kinemage Exercise 7-1

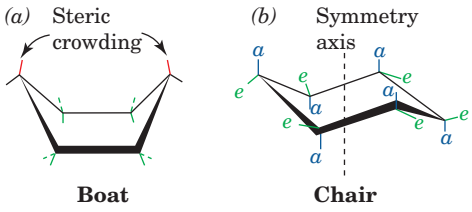



***Figure 11-6*** **Conformations of the cyclohexane ring.** (*a*) In the boat conformation, substituents at the "bow" and "stern" (*red*) are sterically crowded, whereas those along its sides (*green*) are eclipsed. (*b*) In the chair conformation, the substituents that extend parallel to the ring's threefold rotation axis are designated axial [*a*] and those that extend roughly outward from this symmetry axis are designated equatorial [*e*]. The equatorial substituents about the ring are staggered so that they alternately extend above and below the mean plane of the ring.

accurately described as cyclic polyhydroxy hemiacetals or hemiketals.

### b. Sugars Are Conformationally Variable

Hexoses and pentoses may each assume pyranose or furanose forms. The equilibrium composition of a particular monosaccharide depends somewhat on conditions but mostly on the identity of the monosaccharide. For instance, NMR measurements indicate that whereas glucose almost exclusively assumes its pyranose form in aqueous solutions, fructose is 67% pyranose and 33% furanose, and ribose is 75% pyranose and 25% furanose (although in polysaccharides, glucose, fructose, and ribose residues are exclusively in their respective pyranose, furanose, and furanose forms). Although, in principle, hexoses and larger sugars can form rings of seven or more atoms, such rings are rarely observed because of the greater stabilities of the five- and six-membered rings that these sugars can also form. The internal strain of three- and four-membered sugar rings makes them unstable with respect to linear forms.

The use of Haworth formulas may lead to the erroneous impression that furanose and pyranose rings are planar. This cannot be the case, however, because all of the atoms in these rings are tetrahedrally ($sp^3$) hybridized. The pyranose ring, like the cyclohexane ring, may assume a **boat** or a **chair** conformation (Fig. 11-6). The relative stabilities of these various conformations depend on the stereochemical interactions between the substituents on the ring. The boat conformer crowds the substituents on its "bow" and "stern" and eclipses those along its sides, so that in cyclohexane it is ~25 kJ · mol$^{-1}$ less stable than the chair conformer. The ring substituents on the chair conformer (Fig. 11-6*b*) fall into two geometrical classes: the rather close-fitting **axial** groups that extend parallel to the ring's threefold rotational axis and the staggered, and therefore minimally encumbered, **equatorial** groups. Since the axial and equatorial groups on a cyclohexane ring are conformationally interconvertible, a given ring has two alternative chair forms (Fig. 11-7); the one that predominates usually has the lesser crowding among its axial substituents. The conformational situation of a group directly affects its chemical reactivity. For example, equatorial OH groups on pyranoses esterify more readily than do axial OH groups. Note that β-D-glucose is the only D-aldohexose that can simultaneously have all five non-H substituents in the equatorial position (left side of Fig. 11-7). Perhaps this is why glucose is the most abundant naturally occurring monosaccharide. The conformational properties of furanose rings are discussed in Section 29-2Ab in relation to their effects on the conformations of nucleic acids.

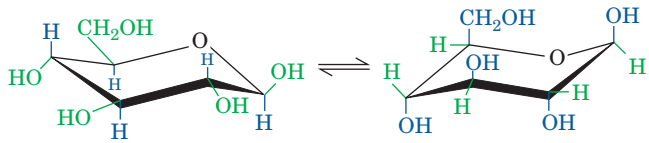


***Figure 11-7*** **The two alternative chair conformations of β-D-glucopyranose.** In the conformation on the left, which predominates, the relatively bulky OH and $CH_2OH$ substituents all occupy equatorial positions, whereas in that on the right (drawn in ball-and-stick form in Fig. 11-5, *right*) they occupy the more crowded axial positions. **See Kinemage Exercise 7-1**

## C. *Sugar Derivatives*

### a. Polysaccharides Are Held Together by Glycosidic Bonds

*The chemistry of monosaccharides is largely that of their hydroxy and carbonyl groups.* For example, in an acid-catalyzed reaction, the anomeric hydroxyl of a sugar reversibly condenses with alcohols to form α- and β-**glycosides** (Greek: *glykys,* sweet) (Fig. 11-8). The bond connecting the anomeric carbon to the acetal oxygen is termed a **glycosidic**

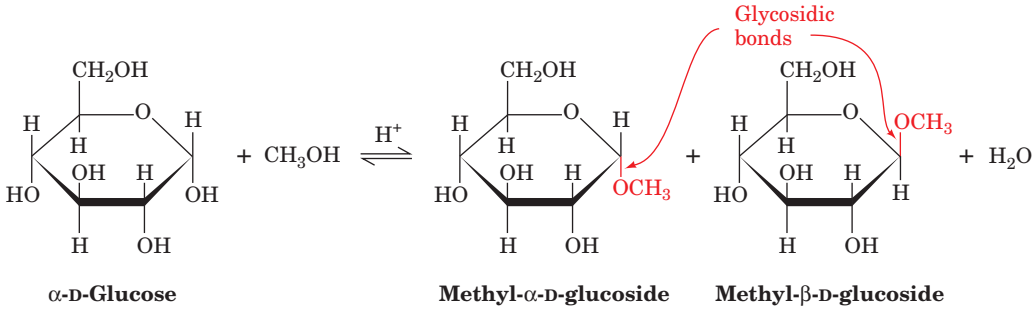


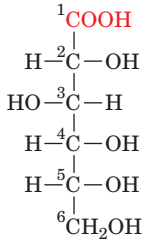

***Figure 11-8*** **The acid-catalyzed condensation of α-D-glucose with methanol to form an anomeric pair of methyl-D-glucosides.**

**bond.** *Polysaccharides are held together by glycosidic bonds between neighboring monosaccharide units.* The glycosidic bond is therefore the carbohydrate analog of the peptide bond in proteins. The bond in a nucleoside linking its ribose residue to its base is also a glycosidic bond (Section 5-1A).

The hydrolysis of glycosidic bonds is catalyzed by enzymes known as **glycosidases** that differ in specificity according to the identity and anomeric configuration of the glycoside but are often rather insensitive to the identity of the alcohol residue. Under basic or neutral conditions and in the absence of glycosidases, however, the glycosidic bond is stable, so glycosides do not undergo mutarotation as do monosaccharides. The methylation of the nonanomeric OH groups of monosaccharides requires more drastic conditions than is required for the formation of methyl glycosides, such as treatment with dimethyl sulfate.

### b. Oxidation–Reduction Reactions

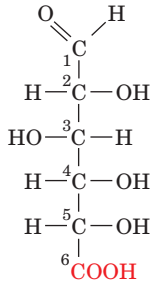
Because the cyclic and linear forms of aldoses and ketoses interconvert so readily, these sugars undergo reactions typical of aldehydes and ketones. Mild oxidation of an aldose, either chemically or enzymatically, results in the conversion of its aldehyde group to a carboxylic acid function, thereby yielding an **aldonic acid** such as **gluconic acid.** Aldonic acids are named by appending the suffix *-onic acid* to the root name of the parent aldose.



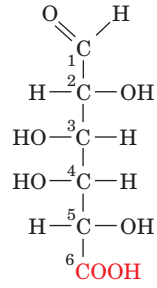
**D-Gluconic acid**

Saccharides bearing anomeric carbon atoms that have not formed glycosides are termed **reducing sugars** because of the facility with which the aldehyde group reduces mild oxidizing agents. A classic test for the presence of a reducing sugar is the reduction of $Ag^+$ in an ammonia solution **(Tollens' reagent)** to yield a metallic silver mirror lining on the inside of the reaction vessel.

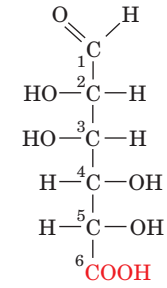
The specific oxidation of the primary alcohol group of aldoses yields **uronic acids,** which are named by appending *-uronic acid* to the root name of the parent aldose. **D-Glucuronic acid, D-galacturonic acid,** and **D-mannuronic acid** are important components of many polysaccharides.



**D-Glucuronic acid**



**D-Galacturonic acid**

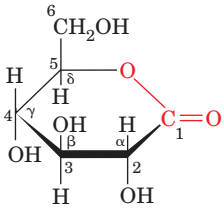


**D-Mannuronic acid**

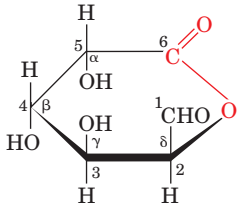
Uronic acids can assume the pyranose, furanose, and linear forms.

Both aldonic and uronic acids have a strong tendency to internally esterify so as to form five- and six-membered lactones (Fig. 11-9). **Ascorbic acid** (**vitamin C,** Fig. 11-10) is a γ-lactone that is synthesized by plants and almost all animals except primates and guinea pigs. Its prolonged deficiency in the diet of humans results in the disease known as **scurvy,** which is caused by the impairment of collagen formation (Section 8-2B). Scurvy generally results from a lack of fresh food. This is because, under physiological conditions, ascorbic acid is reversibly oxidized to **dehydroascorbic acid,** which, in turn, is irreversibly hydrolyzed to the vitamin-inactive **diketogulonic acid** (Fig. 11-10).

Aldoses and ketoses may be reduced under mild conditions, for example, by treatment with $NaBH_4$ to yield acyclic polyhydroxy alcohols known as **alditols,** which are named by appending the suffix *-itol* to the root name of the parent aldose. **Ribitol** is a component of flavin coenzymes (Section 16-2C), and **glycerol** and the cyclic polyhydroxy alcohol ***myo*-inositol** are important lipid components

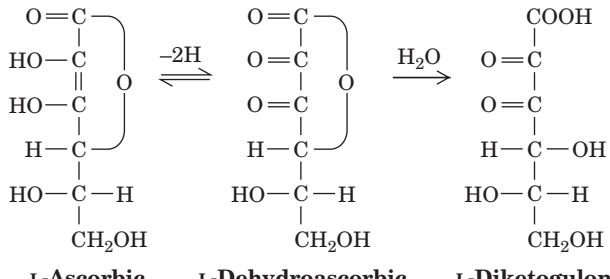


**D-Glucono-δ-lactone**



**D-Glucurono-δ-lactone**

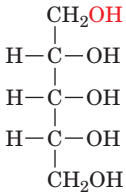
***Figure 11-9*** **D-Glucono-δ-lactone and D-glucurono-δ-lactone are, respectively, the lactones of D-gluconic acid and D-glucuronic acid.** The δ indicates that the O atom closing the lactone ring is also substituent to $C_\delta$.



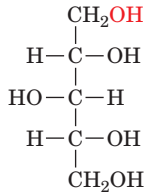
**L-Ascorbic acid** **L-Dehydroascorbic acid** **L-Diketogulonic acid**

***Figure 11-10*** **The reversible oxidation of L-ascorbic acid to L-dehydroascorbic acid.** This is followed by the physiologically irreversible hydrolysis of its lactone ring to form L-diketogulonic acid.

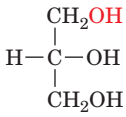
(Section 12-1). **Xylitol** is a sweetener that is used in "sugarless" gum and candies.



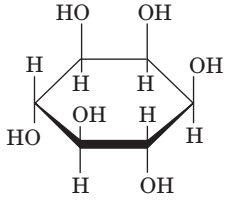
**Ribitol**



**Xylitol**



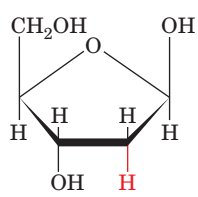
**Glycerol**



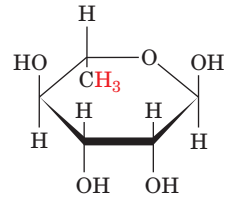
***myo*-Inositol**

### c. Other Biologically Important Sugar Derivatives

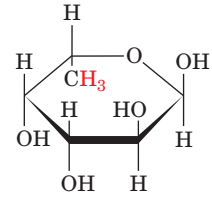
Monosaccharide units in which an OH group is replaced by H are known as **deoxy sugars.** The biologically most important of these is **β-D-2-deoxyribose,** the sugar component of DNA's sugar–phosphate backbone (Section 5-1A). **L-Rhamnose** and **L-fucose** are widely occurring polysaccharide components.



**β-D-2-Deoxyribose**

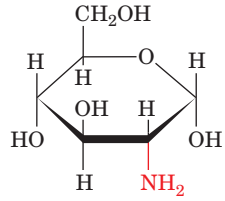


**α-L-Rhamnose (6-deoxy-L-mannose)**

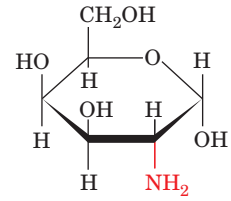


**α-L-Fucose (6-deoxy-L-galactose)**

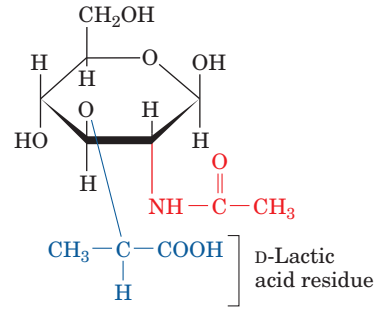
In **amino sugars,** one or more OH groups are replaced by an often acetylated amino group. **D-Glucosamine** and **D-galactosamine** are components of numerous biologically important polysaccharides.



**α-D-Glucosamine (2-amino-2-deoxy-α-D-glucopyranose)**

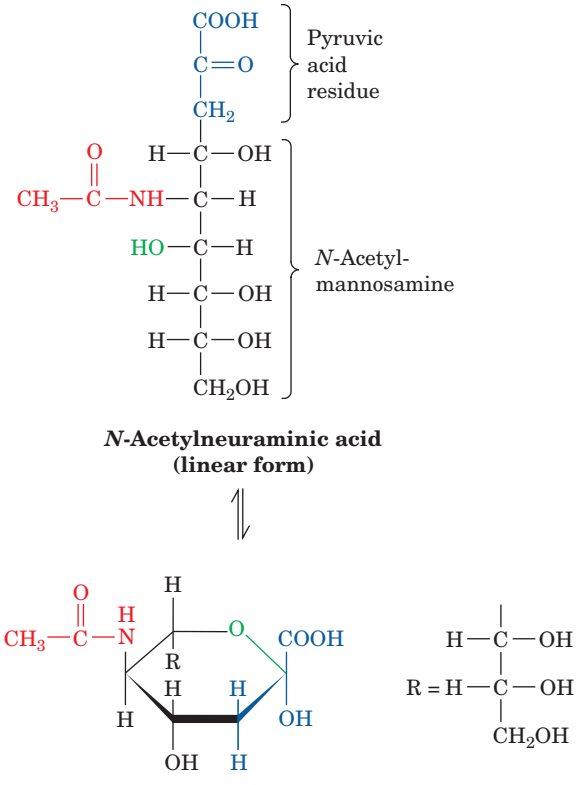


**α-D-Galactosamine (2-amino-2-deoxy-α-D-galactopyranose)**




***N*-Acetylmuramic acid (NAM)**

The amino sugar derivative ***N*-acetylmuramic acid,** which consists of ***N*-acetyl-D-glucosamine** in an ether linkage with **D-lactic acid,** is a prominent component of bacterial cell walls (Section 11-3Ba). *N*-Acetylneuraminic acid, which is derived from *N*-acetylmannosamine and pyruvic acid (Fig. 11-11), is an important constituent of glycoproteins (Section 11-3C) and glycolipids (Section 12-1D). *N*-Acetylneuraminic acid and its derivatives are often referred to as **sialic acids.**




***N*-Acetylneuraminic acid (linear form)**

***N*-Acetylneuraminic acid (pyranose form)**

***Figure 11-11*** ***N*-Acetylneuraminic acid in its linear and pyranose forms.** Note that its pyranose ring incorporates the pyruvic acid residue (*blue*) and part of the mannose moiety.

## 2 POLYSACCHARIDES

Polysaccharides, which are also known as **glycans,** consist of monosaccharides linked together by glycosidic bonds. They are classified as **homopolysaccharides** or **heteropolysaccharides** if they consist of one type or more than one type of monosaccharide residue. Homopolysaccharides may be further classified according to the identity of

their monomeric unit. For example, **glucans** are polymers of glucose, whereas **galactans** are polymers of galactose. Although monosaccharide sequences of heteropolysaccharides can, in principle, be as varied as those of proteins, they are usually composed of only a few types of monosaccharides that alternate in a repetitive sequence.

*Polysaccharides, in contrast to proteins and nucleic acids, form branched as well as linear polymers.* This is because glycosidic linkages can be made to any of the hydroxyls of a monosaccharide. Fortunately for structural biochemists, many polysaccharides are linear and those that branch tend to do so in only a few well-defined ways.

In this section, we discuss the structures of the simplest polysaccharides, the disaccharides, and then consider the structures and properties of the most abundant classes of polysaccharides. We begin by outlining how polysaccharide structures are elucidated.

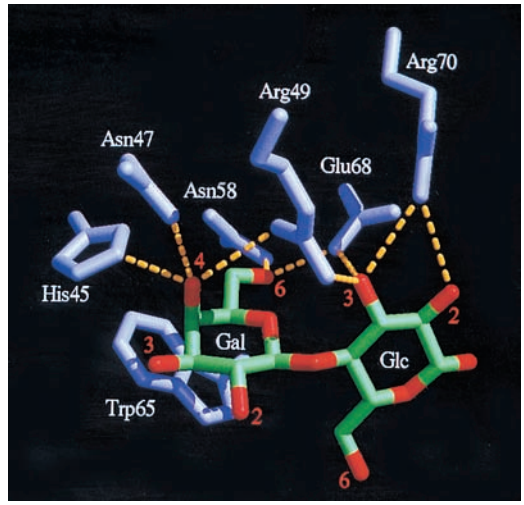



***Figure 11-12*** **Carbohydrate binding by a lectin in the X-ray structure of human galectin-2 in complex with the disaccharide lactose.** This lectin primarily binds β-D-galactose residues. The structure is drawn in stick form with the C and O atoms of lactose's galactose (Gal) and glucose (Glc) residues green and red, and the galectin-2 amino acid side chains violet. Hydrogen bonds between the protein side chains and the sugar residues are represented by dashed yellow lines. [Courtesy of Hakon Leffler, Lund University, Sweden. PDBid 1HLC.]

## A. *Carbohydrate Analysis*

The purification of carbohydrates can, by and large, be effected by chromatographic and electrophoretic procedures similar to those used in protein purification (Sections 6-3 and 6-4), although thin layer chromatography (TLC; Section 6-3Dd) is also widely used. Affinity chromatography (Section 6-3C), using immobilized proteins known as **lectins** (Latin: *legere,* to pick or choose), is a particularly powerful technique in this regard. Lectins are sugar-binding proteins that were discovered in plants but are now known to occur in all organisms, where they participate in a wide variety of signaling, cell–cell recognition, and adhesion processes, as well as in targeting newly synthethesized proteins to specific cellular locations (Section 12-4Cg). Lectins recognize one or more specific monosaccharides with particular linkages to other sugars in oligosaccharides, usually with exquisite specificity. Their protein–carbohydrate interactions typically include multiple hydrogen bonds, which often include bridging water molecules, and the packing of hydrophobic sugar faces against aromatic side chains (Fig. 11-12). Among the best characterized lectins are jack bean **concanavalin A** (Fig. 8-40), which specifically binds α-D-glucose and α-D-mannose residues, and wheat germ **agglutinin** (so named because it causes cells to agglutinate or clump together), which specifically binds β-*N*-acetylmuramic acid and α-*N*-acetylneuraminic acid.

Characterization of an oligosaccharide requires that the identities, anomers, linkages, and order of its component monosaccharides be elucidated. The linkages of the monosaccharides may be determined by **methylation analysis** (also called **permethylation analysis),** a technique pioneered by Norman Haworth in the 1930s. *Methyl ethers not at the anomeric C atom are resistant to acid hydrolysis but glycosidic bonds are not. Consequently, if an oligosaccharide is exhaustively methylated and then hydrolyzed, the free OH groups on the resulting methylated monosaccharides mark the former positions of the glycosidic bonds.* Methylated monosaccharides are often identified by **gas–liquid chromatography** (**GLC;** a technique in which the stationary phase is an inert solid, such as diatomaceous earth, impregnated with a low-volatility liquid, such as silicone oil, and the mobile phase is an inert gas, such as He, into which the sample has been flash evaporated) combined with mass spectrometry (GLC/MS). HPLC techniques may similarly be used. Other mass spectrometric techniques for analyzing nonvolatile substances are discussed in Section 7-1I. Although all aldoses and ketoses with the same number of C atoms are isomers (Figs. 11-1 and 11-2) and hence have identical molecular masses, they have characteristic fragmentation patterns.

The sequence and anomeric configurations of the monosaccharides in an oligosaccharide can be determined through the use of specific **exoglycosidases.** These enzymes specifically hydrolyze their corresponding monosaccharides from the nonreducing ends of oligosaccharides (the ends lacking a free anomeric carbon atom) in a manner analogous to the actions of exopeptidases on proteins (Section 7-1Ab). For example, **β-galactosidase** excises the terminal β anomers of galactose, whereas **α-mannosidase** does so with the α anomers of mannose. Some of these exoglycosidases also exhibit specificity for the **aglycone,** the sugar chains to which the monosaccharide to be excised (the **glycone**) is linked. Through the use of mass spectrometry, the sequence of a polysaccharide may be deduced from the mass decrements generated by exoglycosidases. The use of **endoglycosidases** (hydrolases that cleave glycosidic bonds between nonterminal sugar residues) of varying specificities can also supply useful sequence information. The proton and $^{13}C$ NMR spectra of oligosaccharides can provide the complete sequence of an oligosaccharide if sufficient material is available. Moreover, two-dimensional

NMR techniques (Section 8-3Ac) can reveal oligosaccharide structures (e.g., see Section 11-2Eb).

### B. *Disaccharides*

We begin our studies of polysaccharides by considering disaccharides (Fig. 11-13). **Sucrose,** the most abundant disaccharide, occurs throughout the plant kingdom and is familiar to us as common table sugar. Its structure (Fig. 11-13) was established by methylation analysis as described above and was later confirmed by its X-ray structure. To name a polysaccharide systematically, one must specify its component monosaccharides, their ring types, their anomeric forms, and how they are linked together. Sucrose is therefore *O*-α-D-glucopyranosyl-(1 → 2)-β-D-fructofuranoside, where the symbol (1 → 2) indicates that the glycosidic bond links C1 of the glucose residue to C2 of the fructose residue. Note that since these two positions are the anomeric carbon atoms of their respective monosaccharides, sucrose is not a reducing sugar (as the suffix *-ide* implies).

The hydrolysis of sucrose to D-glucose and D-fructose is accompanied by a change in optical rotation from *dextro* to *levo.* Consequently, hydrolyzed sucrose is sometimes called **invert sugar** and the enzyme that catalyzes this process, **β-D-fructofuranosidase,** is archaically named **invertase.**

**Lactose** [*O*-β-D-galactopyranosyl-(1 → 4)-D-glucopyranose] or milk sugar (Fig. 11-13) occurs naturally only in milk, where its concentration ranges from 0 to 7% depending on the species. The free anomeric carbon of its glucose residue makes lactose a reducing sugar.

Infants normally express the intestinal enzyme **β-D-galactosidase** or **lactase** that catalyzes the hydrolysis of lactose to its component monosaccharides for absorption into the bloodstream. Many adults, however, including most Africans and almost all Asians, have a low level of this enzyme (as do most adult mammals, since they normally do not encounter milk). Consequently, much of the lactose in any milk they drink moves through their digestive tract to the colon, where its bacterial fermentation produces large quantities of $CO_2$, $H_2$, and irritating organic acids. This results in an embarrassing and often painful digestive upset termed **lactose intolerance.** Perhaps this is why Chinese cuisine, which is noted for the wide variety of foodstuffs it employs, is devoid of milk products. However, adult members of populations with a tradition of herding cattle, mainly northern Europeans and certain African groups, continue expressing the lactase gene and hence can drink milk without a problem. Modern food technology has come to the aid of milk lovers who develop lactose intolerance: Milk products in which the lactose has been hydrolyzed enzymatically and lactase-containing pills are now widely available.

There are several common glucosyl–glucose disaccharides. These include **maltose** [*O*-α-D-glucopyranosyl-(1 → 4)-D-glucopyranose], an enzymatic hydrolysis product of starch; **isomaltose,** its α(1 → 6) isomer; and **cellobiose,** its β(1 → 4) isomer, the repeating disaccharide of cellulose. Only a few tri- or higher oligosaccharides occur in significant natural abundance. Not surprisingly, they all occur in plants.

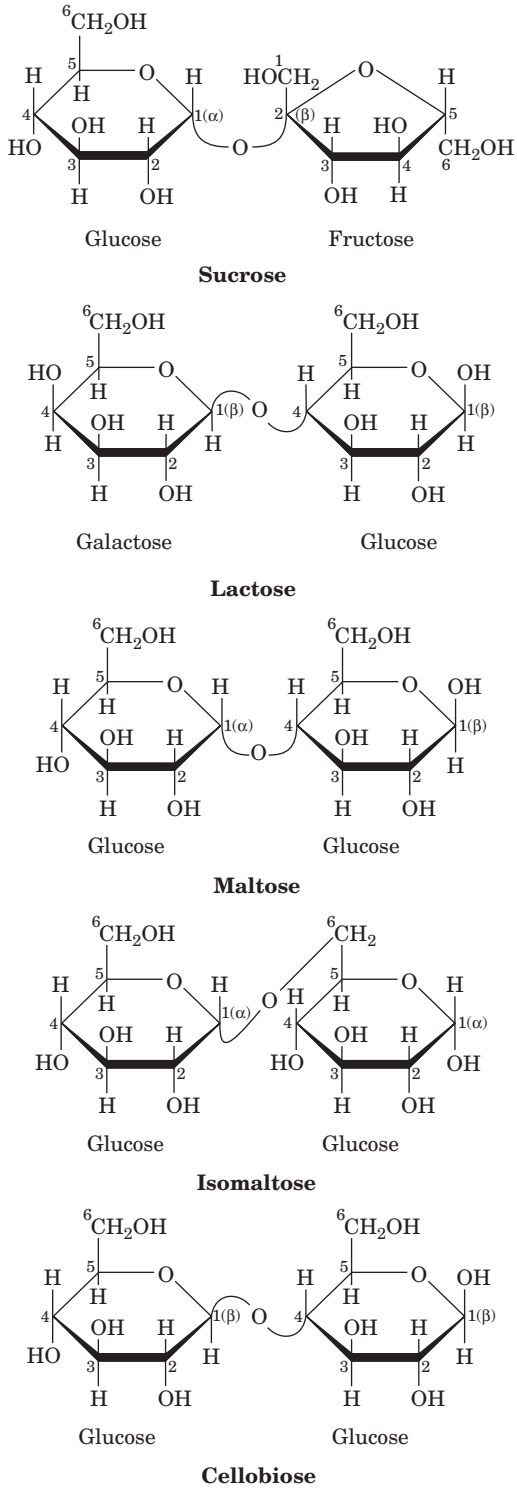



***Figure 11-13*** **Several common disaccharides.** **See Kinemage Exercise 7-2**

### C. *Structural Polysaccharides: Cellulose and Chitin*

Plants have rigid cell walls (Fig. 1-9) that, in order to maintain their shapes, must be able to withstand osmotic pressure differences between the extracellular and intracellular spaces of up to 20 atm. In large plants, such as trees, the cell

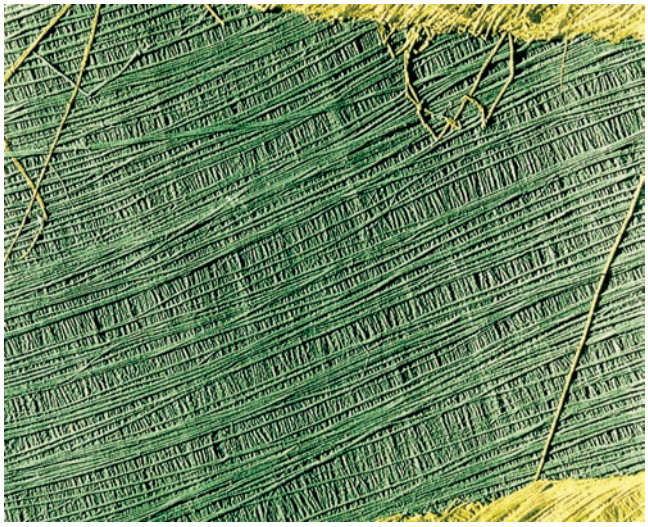
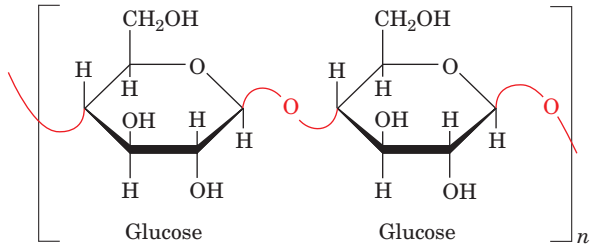


***Figure 11-14*** **Electron micrograph of the cellulose fibers in the cell wall of the alga *Chaetomorpha melagonium*.** Note that the cell wall consists of layers of parallel fibers. [Biophoto Associates/Photo Researchers.]




**Cellulose**

***Figure 11-15*** **The primary structure of cellulose.** Here *n* may be several thousand.

walls also have a load-bearing function. Cellulose, the primary structural component of plant cell walls (Fig. 11-14), accounts for over half of the carbon in the biosphere: $\sim 10^{15}$ kg of cellulose are estimated to be synthesized and degraded annually. Although cellulose is predominantly of vegetable origin, it also occurs in the stiff outer mantles of marine invertebrates known as **tunicates** (urochordates; Fig. 1-11).

The primary structure of cellulose was determined through methylation analysis. Cellulose is a linear polymer of up to 15,000 D-glucose residues (a glucan) linked by $\beta(1 \rightarrow 4)$ glycosidic bonds (Fig. 11-15). As is generally true of large polysaccharides, it has no defined size since, in contrast to proteins and nucleic acids, there is no genetically determined template that directs its synthesis.

X-ray studies of cellulose fibers led Anatole Sarko to tentatively propose the model diagrammed in Fig. 11-16. This highly cohesive, hydrogen bonded structure gives cellulose fibers exceptional strength and makes them water insoluble despite their hydrophilicity.

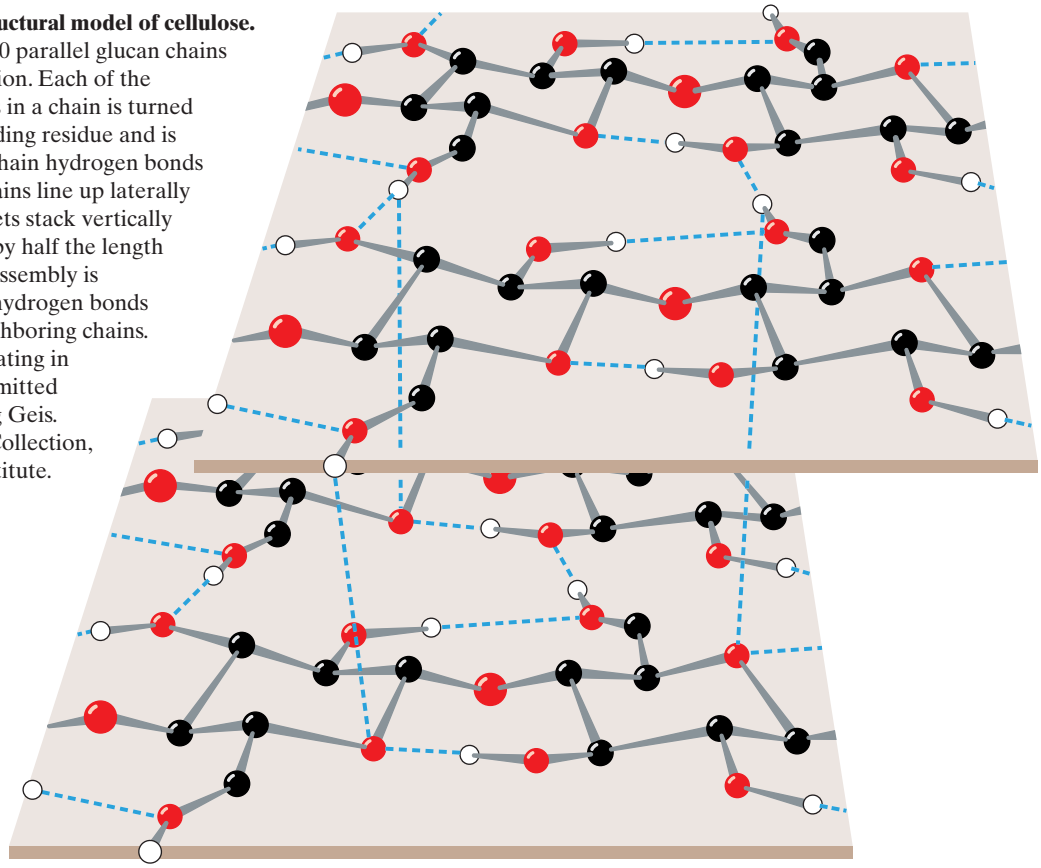
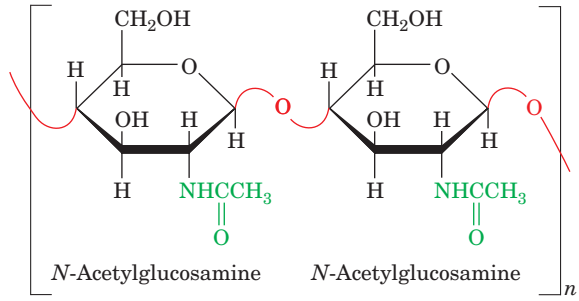


***Figure 11-16*** **Proposed structural model of cellulose.** Cellulose fibers consist of ~40 parallel glucan chains arranged in an extended fashion. Each of the $\beta(1 \rightarrow 4)$-linked glucose units in a chain is turned 180° with respect to its preceding residue and is held in this position by intrachain hydrogen bonds (*dashed lines*). The glucan chains line up laterally to form sheets, and these sheets stack vertically such that they are staggered by half the length of a glucose unit. The entire assembly is stabilized by intermolecular hydrogen bonds between glucose units of neighboring chains. Hydrogen atoms not participating in hydrogen bonds have been omitted for clarity. [Illustration, Irving Geis. Image from the Irving Geis Collection, Howard Hughes Medical Institute. Reprinted with permission.]

In plant cell walls, the cellulose fibers are embedded in and cross-linked by a matrix of several polysaccharides that are composed of glucose as well as other monosaccharides. In wood, this cementing matrix also contains a large proportion of **lignin,** a plasticlike phenolic polymer. One has only to watch a tall tree in a high wind to realize the enormous strength of plant cell walls. In engineering terms, they are "composite materials," as is concrete reinforced by steel rods. Composite materials can withstand large stresses because the matrix evenly distributes the stresses among the reinforcing elements.

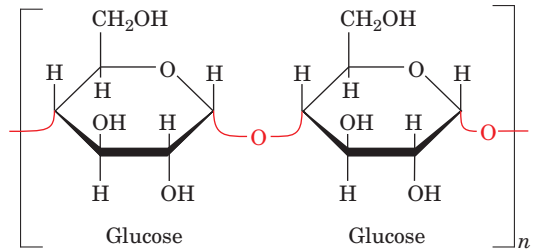
Although vertebrates themselves do not possess an enzyme capable of hydrolyzing the β(1 → 4) linkages of cellulose, the digestive tracts of herbivores contain symbiotic microorganisms that secrete a series of enzymes, collectively known as **cellulase,** that do so. The same is true of termites. Nevertheless, the degradation of cellulose is a slow process because its tightly packed and hydrogen bonded glucan chains are not easily accessible to cellulase and do not separate readily even after many of their glycosidic bonds have been hydrolyzed. The digestion of fibrous plants such as grass by herbivores is therefore a more complex and time-consuming process than is the digestion of meat by carnivores (cows, e.g., have multichambered stomachs and must chew their cud). Similarly, the decay of dead plants by fungi, bacteria, and other organisms, and the consumption of wooden houses by termites, often takes years.

**Chitin** is the principal structural component of the exoskeletons of invertebrates such as crustaceans, insects, and spiders and is also a major cell wall constituent of most fungi and many algae. It is estimated that ~$10^{14}$ kg of chitin are produced annually, most of it in the oceans, and therefore that it is almost as abundant as is cellulose. Chitin is a homopolymer of β(1 → 4)-linked *N*-acetyl-D-glucosamine residue (Fig. 11-17). It differs chemically from cellulose only in that each C2-OH group is replaced by an acetamido function. X-ray analysis indicates that chitin and cellulose have similar structures.

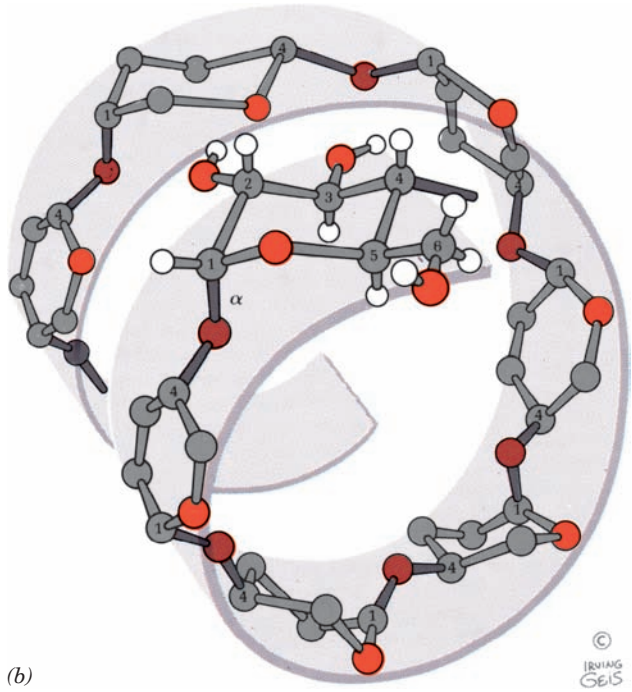



**Chitin**

***Figure 11-17*** **The primary structure of chitin.** Chitin is a β(1 → 4)-linked homopolymer of *N*-acetyl-D-glucosamine.




*(a)* **α-Amylose**




*(b)*

***Figure 11-18*** **α-Amylose.** (*a*) The D-glucose residues of α-amylose are linked by α(1 → 4) bonds (*red*). Here *n* is several thousand. (*b*) This regularly repeating polymer forms a left-handed helix with ~6 glucose residues per turn. Note the great differences in structure and properties that result from changing α-amylose's α(1 → 4) linkages to the β(1 → 4) linkages of cellulose (Fig. 11-16). [Illustration, Irving Geis. Image from the Irving Geis Collection, Howard Hughes Medical Institute. Reprinted with permission.]

## D. *Storage Polysaccharides: Starch and Glycogen*

### a. Starch Is a Food Reserve in Plants and a Major Nutrient for Animals

Starch is a mixture of glucans that plants synthesize as their principal food reserve. It is deposited in the cytoplasm of plant cells as insoluble granules composed of **α-amylose** and **amylopectin.** α-Amylose is a linear polymer of several thousand glucose residues linked by α(1 → 4) bonds (Fig. 11-18*a*). Note that although α-amylose is an isomer of cellulose, it has very different structural properties. This is because cellulose's β-glycosidic linkages cause each successive glucose residue to flip 180° with respect to the preceding residue, so that the polymer assumes an easily packed, fully extended conformation (Fig. 11-16). In contrast, α-amylose's α-glycosidic bonds cause it to adopt an irregularly aggregating helically coiled conformation (Fig. 11-18*b*).

Amylopectin consists mainly of α(1 → 4)-linked glucose residues but is a branched molecule with α(1 → 6) branch

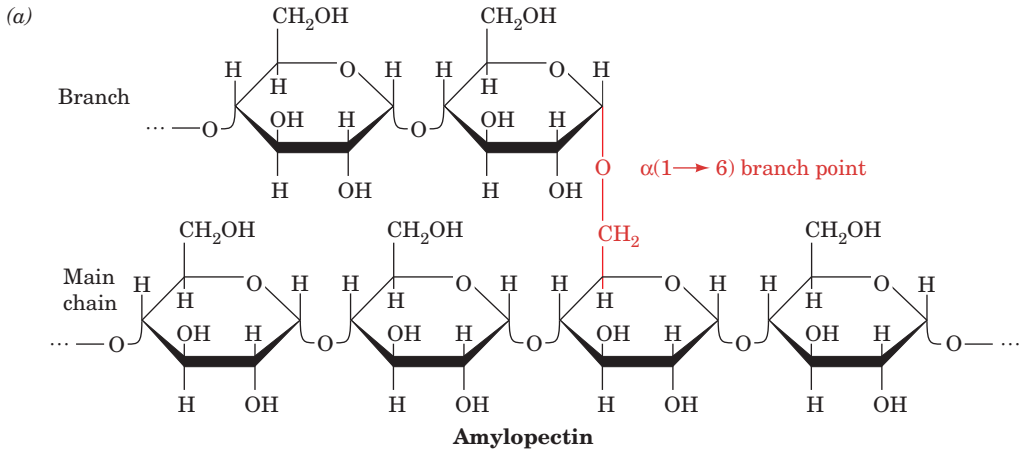



***Figure 11-19*** **Amylopectin.** (*a*) Its primary structure near one of its α(1 → 6) branch points (*red*). (*b*) Its bushlike structure with glucose residues at branch points indicated in red. The actual distance between branch points averages 24 to 30 glucose residues. Glycogen has a similar structure but is branched every 8 to 14 residues.

points every 24 to 30 glucose residues on average (Fig. 11-19). Amylopectin molecules contain up to $10^6$ glucose residues, which makes them among the largest molecules occurring in nature. The storage of glucose as starch greatly reduces the large intracellular osmotic pressures that would result from its storage in monomeric form because osmotic pressure is proportional to the number of solute molecules in a given volume.

### b. Starch Digestion Occurs in Stages

The digestion of starch, the main carbohydrate source in the human diet, begins in the mouth. Saliva contains **α-amylase,** which randomly hydrolyzes all the α(1 → 4) glucosidic bonds of starch except its outermost bonds and those next to branches. By the time thoroughly chewed food reaches the stomach, where the acidity inactivates α-amylase, the average chain length of starch has been reduced from several thousand to fewer than eight glucose units. Starch digestion continues in the small intestine under the influence of pancreatic α-amylase, which is similar to the salivary enzyme. This enzyme degrades starch to a mixture of the disaccharide maltose, the trisaccharide **maltotriose,** which contains three α(1 → 4)-linked glucose residues, and oligosaccharides known as **dextrins** that contain the α(1 → 6) branches. These oligosaccharides are hydrolyzed to their component monosaccharides by specific enzymes contained in the brush border membranes of the intestinal mucosa: an **α-glucosidase,** which removes one glucose residue at a time from oligosaccharides, an **α-dextrinase** or **debranching enzyme,** which hydrolyzes α(1 → 6) and α(1 → 4) bonds, a **sucrase,** and, at least in infants, a lactase. The resulting monosaccharides are absorbed by the intestine and transported to the bloodstream (Section 20-4A).

### c. Glycogen Is "Animal Starch"

Glycogen, the storage polysaccharide of animals, is present in all cells but is most prevalent in skeletal muscle and liver, where it occurs as cytoplasmic granules (Fig. 11-20). The primary structure of glycogen resembles that of amylopectin, but glycogen is more highly branched, with branch points occurring every 8 to 14 glucose residues. Glycogen's degree of polymerization is nevertheless similar to that of amylopectin. In the cell, glycogen is degraded for metabolic use by **glycogen phosphorylase,** which phosphorolytically cleaves glycogen's α(1 → 4) bonds sequentially inward from its nonreducing ends to yield **glucose-1-phosphate.** Glycogen's highly branched structure, which has many nonreducing ends, permits the rapid mobilization of glucose in times of metabolic need. The α(1 → 6) branches of glycogen are cleaved by a debranching enzyme. These enzymes play an important role in glucose metabolism and are discussed further in Section 18-1.

## E. *Glycosaminoglycans*

The extracellular spaces, particularly those of connective tissues such as cartilage, tendon, skin, and blood vessel walls, consist of collagen and elastin fibers (Section 8-2B)